Title: From vineyard to winery: a source map of microbial diversity driving wine fermentation

Authors: Peter Morrison-Whittle[a], and Matthew R Goddard[a,b]

Affiliations:

- [a] The School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.
- [b] School of Life Sciences, and Lincoln Institute for Agri-Food Technology, University of Lincoln, Lincoln, LN6 7DL, United Kingdom

Email addresses: pmor072@aucklanduni.ac.nz, mgoddard@lincoln.ac.uk

Running title: The origins of microbial diversity in wine.

Type of article: Research paper

Corresponding Author: Peter Morrison-Whittle, The School of Biological Sciences, The University of Auckland, Private Bag 92019, Auckland 1142, New Zealand.

pmor072@aucklanduni.ac.nz

Conflict of interest: We are not aware of any conflict of interest in carrying out this study.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1111/1462-2920.13960

Originality-Significance Statement: We show that ~40% juice/ferment fungal communities overlap with vineyard fungi communities, which includes the most abundant species in both juice and ferment. While this may seem obvious, this has not been robustly shown before. As far as we are aware, this is the first study to analytically quantify the connectivity of microbial communities in native habitats to managed agricultural ecosystems. For viticulture and wine, this is also the first demonstration that significant regional differences in microbial communities in commercial juice correlate with communities in local ecosystems. We also show that while the fungal communities present in juice show regional delineations and initially resemble those found on grapes, surprisingly during fermentation these communities resemble those present on the bark of local grapevines. In addition, these data show that ~30% of species in ferments may also be found in local native forests.

From vineyard to winery: a source map of microbial diversity driving wine fermentation

Peter Morrison-Whittle[a] and Matthew R Goddard[a,b]

[a] The School of Biological Sciences, The University of Auckland, New Zealand

and

[b] The School of Life Sciences, and Lincoln Institute for Agri-Food Technology, The University of Lincoln, United Kingdom

Summary

Humans have been making wine for thousands of years and microorganisms play an integral part in this process as they not only drive fermentation, but also significantly influence the flavour, aroma, and quality of finished wines. Since fruits are ephemeral, they cannot

comprise a permanent microbial habitat; thus, an age-old unanswered question concerns the origin of fruit and ferment associated microbes. Here we use next-generation sequencing approaches to examine and quantify the roles of native forests, vineyard soil, bark, and fruit habitats as sources of fungal diversity in ferments. We show that aspects of microbial communities in harvested juice and ferments vary significantly across regions, and that while vineyard fungi account for ~40% of the source of this diversity, uncultivated ecosystems outside of vineyards also prove a significant source. We also show that while communities in harvested juice resemble those found on grapes, these increasingly resemble fungi present on vine bark as the ferment proceeds.

Introduction

Many different species of fungi and bacteria are found naturally associated with fruits and their spontaneous fermentation (Zott et al., 2010; Barata et al., 2012; Bokulich et al., 2014; Taylor et al., 2014), and these play a crucial role influencing the flavour, aroma, quality, and thus value of commercial wines and other fermented beverages (Pretorius, 2000; Ciani et al., 2010; Knight, 2015). Despite being the focus of commercial and scientific interest, we know little about the spatial ecology of most microbial species generally associated with vines and wines other than a handful of pathogens (Goddard and Greig, 2015). Grape bunches, the primary substrate of winemaking, are perhaps the most obvious potential source of microbial diversity within ferments. The problem is, grapes, like all fruits, are ephemeral and thus cannot be a stable habitat for microbes; all fruit associated microbes

must have migrated from elsewhere. This gives rise to a long-unanswered question – what are the origins of the microbial species associated with grapes and their ferments?

To date, much work has focused on just one member of this community, the crabtree positive *S. cerevisiae,* due to both its importance in fermentation and wine quality, and as a research model (Ciani et al., 2002, 2004; Goddard, 2008; Goddard and Greig, 2015). Increasing genetic evidence now links populations and specific strains of *S. cerevisiae* found in spontaneous ferments in wineries to those also recovered from vineyards from which the fruit derived, as well as surrounding unmanaged forest niches (Goddard et al., 2010; Knight and Goddard, 2015; Börlin et al., 2016). However, *Saccharomyces* species only constitute a very small fraction of the microbial community associated with fruit prior to the mid-stage of fermentation (Goddard, 2008). An array of other Ascomycete species contribute to ferments, interact with *Saccharomyces* species, and affect wine aroma, flavour, and other properties, in various positive and negative ways (Ciani et al., 2010; Jolly et al., 2014). The origins of these ferment fungi, which constitute the bulk of ferment diversity, remains unexplained.

Recent microbial surveys of vineyards – arguably the most likely source of ferment diversity – revealed diverse fungal communities whose structure differs at scales of hundreds to thousands of kilometres (Gayevskiy and Goddard, 2011; Taylor et al., 2014; Bokulich et al., 2014; Morrison-Whittle and Goddard, 2015; Pinto et al., 2015). Given the limited but increasing evidence for a microbial aspect to *terroir,* which posits that regional flavours and aromas in wine may be in part due to regionally structured microbial diversity, then uncovering the relationship between ferment diversity and nearby managed or unmanaged

ecosystems represents a valuable target of research. We are aware of no study that has quantified the similarities and connections of fruit microbial communities to surrounding managed and unmanaged habitats, and the resulting spontaneous ferments of fruit from the same area.

Here we test the null-hypothesis predicting there is no connection between fungal communities in fresh pressed grape-juice and their resulting spontaneous fermentation to surrounding managed and natural habitats. To do this, we sampled communities from 36 vineyards and wineries across six New Zealand wine growing regions, as well as 36 nearby native forests. We employed next-generation DNA sequencing of 26S amplicon libraries produced from these samples to estimate fungal communities present in: bark, soil, and grape-bunches within vineyards; soil and fruit samples within unmanaged native forests; and commercially pressed grape juice collected from wineries and their resulting spontaneous ferments.

We initially evaluate patterns of diversity across winery and terrestrial ecosystems at a national, regional, and vineyard level and quantify the microbial links between sampled habitats. We then go on to test whether significant regional differences exist between fungal communities in spontaneous ferments, and then go on to test hypotheses regarding the origin of microbes that spontaneously ferment wine grapes. This study not only explores whether there is a connection between microbial communities in these various habitats, but also quantifies these connections and thus estimates the relative contributions of fungal communities in various environmental habitats to those that conduct fermentation in the

winery. Using this information, we directly address the classic question concerning the origin of these microbes.

Results

In quantifying the similarities between fungal communities in different habitats, we define three possible types of differences among communities: 1 – absolute species richness: differences in numbers of species; 2 - relative species richness: differences in types of species; and 3 - community composition: differential abundances of species. The difference between these is important and using these we may evaluate the nature of the similarity of communities in various habitats. (see Figure 1).

Fungal diversity in juices and ferments

We first examined fungal communities in 33 commercially prepared juices, and from day five of the spontaneous fermentation of these same juices (hereafter designated 'ferment' samples). DNA sequencing of the 26S D1/D2 rDNA revealed a total of 154 molecular operational taxonomic units (which we simply refer to species herein) across all juice and ferment samples. Ascomycota comprised 74% of species, with most of the remaining species identified as Basidiomycota or unclassified (13% for both). In terms of abundances, Ascomycota was the most abundant phyla comprising 92.1% of all sequences, 7.5% of the remaining sequences were unclassified, with Basidiomycete species observed very rarely at 0.4%. Overall, we identified 2 phyla, 11 classes, 14 orders, 23 families, and 34 genera, and

154 species in these juice and ferment habitats. When juice and ferment communities from the same winery were combined, we found an average of 18.4 (±1.3) species per winery. We observed a pronounced shift in fungal diversity from freshly pressed juice to ferment samples across different measures of biodiversity, with members of the *Saccharomycetaceae* family becoming the most abundant taxa (see Figure 2). This shift in diversity is consistent with the fermentative actions of *Saccharomyces* species as these manipulate the environment and competitively exclude non-*Saccharomyces* species (Goddard, 2008). Previous population genetics work with *Saccharomyces cerevisiae* derived from these very same ferments shows that these *S. cerevisiae* populations were largely derived from the surrounding environment (Knight and Goddard, 2014); however, this does not preclude the possibility that some fraction of the strains may have been sourced from 'cellar' populations, though we have no data supporting this hypothesis for these samples.

Regional differences in harvested fruit and ferment communities

Previous studies have shown microbial communities associated with vineyard soil, vine bark, and ripe fruit differ across regions in NZ, including samples from the very same vineyards we study here (Taylor et al., 2014; Morrison-Whittle and Goddard, 2015). Bokulich et al. (2014) show similar patterns in juice communities in California, and Pinto et al. (2015) in Portugal. Given these patterns, it is possible the nature of any connections between environmental and harvested fruit and ferment communities may also vary between regions. We first tested for regional signals in communities, and found there was no effect of region in either juice or ferment in terms of absolute species numbers (Juice: $F_{5,29} = 1.25$, $P = 0.3121$; ferment: $F_{5,27} = 1.87$, $P = 0.1325$). The types of species were significantly different by region

in juice (R^2 = 0.216, P < 0.0001) but not ferment (R^2 = 0.154, P = 0.4954) communities. However, the relative abundances of species in both juice (R^2= 0.261, P < 0.0001) and ferment communities (R^2 = 0.335, P = 0.0044) significantly differed among regions (Table S1). Unsurprisingly, *Saccharomyces* was the most abundant species in ferment, and the breakdown of juice and ferment communities is shown in Figure 2.

Additive diversity partitioning and hierarchical null model testing analyses provide another approach to test whether the species discovered are equally distributed among regions (Crist et al., 2003). We found that the average species richness for both juice and ferment communities were significantly lower than those predicted under random models (see Table 1), indicating that regions tended to have fewer species than expected if they were randomly distributed across NZ. Species turnover at the national level (β_2-diversity) was 2-fold and 3-fold greater than species turnover within regions (β_1-diversity) for both juice and ferment communities, indicating that individual communities were more similar to communities within their own region than to those from other regions (see Figure 3). Both analytical approaches reveal a signal for region-specific differential fungal communities in juices, and to a limited extent in spontaneous ferments in different areas of New Zealand.

On the origin of winery fungi

To evaluate the connection between harvested fruit and ferment communities to those in surrounding environmental habitats, we compared the communities in native forest, vineyard bark, soil, and fruit to those in juice and ferments. Firstly, we evaluated the similarities of species presence and absence in juice and ferment samples with both

combined native and managed environmental samples, and this revealed a substantial overlap of species in ferments with juice as well as both native and managed ecosystems; see Figure 4. When soil, bark, and fruit sub-compartments in managed and native ecosystems were analysed and compared to juice communities, each vineyard habitat (bark, fruit, and soil) shared the same total number of species with juice (36 of the 127 in juice), representing 28.3% of all juice species. Native fruit and soil communities had 32 and 28 species in common with juice communities (representing 25.2% and 22.0% of all juice species) respectively. Similarly, when soil, bark, and fruit sub-compartments in managed and native ecosystems were analysed and compared to species present in ferment communities, then vine bark communities showed the greatest overlap with ferment communities, with 17 of the 56 (30.4%) species in common. Vineyard soil communities had 13 (23.2%) species overlapping with ferments, fruit 10 (17.9%), native soil 9 (16.1%), and finally, native fruit contained 8 (14.3%) species that overlapped with ferment communities. Overall, many species found in either juice or ferment samples were observed in more than one vineyard habitat. Around 14.6% of harvested fruit and ferment species were only found in one vineyard habitat, 12.3% of species in juice were found in two habitats, and 13.0% of species were found in all three. Figure 4 describes the number of similar species in each compartment individually, but if the juice and ferment communities are combined to constitute a 'winemaking habitat', then 64 of the 154 total species (41.6%) were also identified in vineyard habitats (fruit, bark, soil).

Figure 5 shows the relative abundance of the most abundant taxa in juice and ferment samples, as well as their relative frequencies within grape, bark, and soil communities. Similarly, of all species identified in winemaking habitats, 49 (31%) were also identified in

native forest habitats; of these, only 3 species in ferments were found only in native forest habitats.

While the types of species in ferment communities did not significantly differ by region, the types of species in juice did. We then analysed the similarity of juice communities to environmental habitats taking into account the regional differences shown earlier. When juice communities are compared to soil, bark, and fruit communities from the same vineyard from which they were harvested, it appears juice communities have the greatest overlap with fruit communities (see Figure 6). However, as ferments progress, community similarity changes to more greatly reflect bark communities in the respective vineyards (see Figure 6). This result is in-line with the overall analyses reported above, showing the greatest similarity between ferment and bark communities. On average, 34.1% of species found in any one ferment were also found in the vineyard from which the fruit derived.

However, while a considerable number of harvested juice and ferment fungi have their origins potentially accounted for in vineyard and native forest habitats, this still leaves 87 species (56.5% of all juice and ferment species) unaccounted for. While over 50% of species types are not accounted for in vineyard habitats, these species only account for 14.5% in terms of abundance. The potential sources for these species are not clear.

Discussion

We found substantial evidence to reject the null hypotheses of no overlap between winery and environmental fungal communities. We show that ~40% juice/ferment fungal communities overlap with vineyard fungal communities, which includes the most abundant species in both juice and ferment. While this may seem obvious, this has not been robustly shown before. We also show that while the fungal communities present in juice show regional delineations and initially resemble those found on grapes, over the course of fermentation these communities increasingly resemble those present on the bark of local grapevines. In addition, these data show that ~30% of species in ferments may also be found in local native forests. As far as we are aware, this is the first study to analytically quantify the overlap of microbial communities in native habitats to managed agricultural ecosystems. For viticulture and wine, this is also the first demonstration that significant regional differences in microbial communities in commercial juice correlate with communities in local ecosystems.

We show that fungal communities in freshly harvested juice significantly differ across regional scales, and this is in line with similar findings of Bokulich et al. (2014), Pinto et al. (2015), and Taylor et al. (2014). These results also fall broadly in line with other microbial biogeographic studies that have detected non-random patterns of microbial diversity within vineyards (Gayevskiy and Goddard, 2011; Morrison-Whittle and Goddard, 2015) and other terrestrial environments (Martiny et al., 2006; Hanson et al., 2012; Nemergut et al., 2013). The connection described here between vineyard and harvested juice and ferment diversity supports observational reports where certain fermentative species (*Saccharomyces* spp.)

and grapevine pathogens have been reported in both vineyard and winery environments (Mortimer and Polsinelli, 1999; Barata et al., 2008, 2012). Further, the quantitative community ecology connections described here are directly in line with population genetic analyses showing the same genotypes of *S. cerevisiae* are present in the local environment and ferments (Goddard et al., 2010; Knight and Goddard, 2014).

The most obvious explanation for finding regionally distinct juice communities is that these are a function of microbial diversity in local managed and native ecosystems. It is interesting to note that regional differences in community diversity become weaker as the ferment proceeds. This is most likely due to the collapse of microbial diversity during fermentation as *Saccharomyces* displaces other species (Goddard, 2008). However, at the population genetic level, the genotypes (strains) of *S. cerevisiae* retain signals for genetic differentiation between regions (Knight and Goddard, 2014), and these may impart different chemical signatures to wines (Knight et al., 2015). In addition, the various 'non-*Saccharomyces*' yeasts in juices are known to significantly contribute to wine aroma and flavour (Ciani et al., 2010; Jolly et al., 2014), and may also interact with *Saccharomyces* (Anfang et al., 2009). Indeed, this is in line with reports that regional microbiomes of ferments correlate with wine metabolomes (Bokulich et al., 2016). Thus, the potential for the combination of regionally distinct non-*Saccharomyces* communities and *Saccharomyces* strains to influence wine aroma and flavour means the door for a microbial aspect to terroir remains open.

Notably, we also show a number of species present in juice and ferments that are not found in any sampled habitat. One obvious possibility is that a fraction of ferment diversity is seeded from the processing facilities of the wineries themselves, or vectored in by human

workers. As there is evidence that some *Saccharomyces* strains may inhabit wineries between vintages (Ciani et al., 2004; Mercado et al., 2007; Santamaría et al., 2008; Blanco et al., 2011), it is not unreasonable to suggest that wineries may also harbour other fungal species. One recent study shows that an assortment of different species may be found on winery equipment, but that under standard cleaning conditions these communities fluctuate greatly, and do not comprise species typically found in fermentation (Bokulich et al., 2013). Another possibility, as the species unaccounted for tend to be the rarer ones, is that their absence in our environmental samples is a result of insufficient sampling in vineyard and native habitats. Sampling directly from soil, bark and vines may well miss fungal species that are patchily distributed in vineyards, as producing juice from harvested grapes effectively homogenises microbial communities present on very large numbers of grapes.

It is important to note that finding overlapping diversity between two environments does not constitute empirical demonstration of source-sink relationships between those two communities: such communities may simply be abundant in multiple habitats. Empirically demonstrating a connection would require artificial seeding and recovery of molecularly-labelled organisms. The connection revealed between different vineyard habitats and harvested juice and ferments in wineries is not surprising considering the scale and process of mechanical harvesting. The fact that fruit, harvest juice and ferments are not permanent microbial habitats, combined with the flow of material from vineyard to winery means the common-sense interpretation is that microbes flow in the same direction. Additionally, it is also possible that microbial communities in different vineyard habitats may be indirectly connected to winery habitats through animal vectors. Humans are likely candidates for

connecting vineyards to wineries, given the overlap in labour and tool use. Insects could also connect vineyards to wineries as they are not only known to carry ascomycete fungi (Stefanini et al., 2012), but appear to have evolved mutualistic relationships with some species, and this could explain why fungi produce olfactory attractants during fermentation in the first place (Buser et al., 2014).

These results confirm the contribution of vineyard communities to ferment fungal diversity, but challenges the notion that fruit-associated fungi are the most pertinent to elucidating ferment/wine ecology. Our data indicate a substantial fraction of ferment diversity overlaps with diversity in bark and other habitats. Even though vineyard soil and bark material appear to contribute little in the way of biomass to commercial ferments, it appears that microbes from these habitats constitute a disproportionate fraction of ferment diversity. Moreover, the contributions of species from uncultivated habitats appears substantial. This suggests we need to better understand how different agricultural and native ecosystem management methods might affect these microbial communities and thus their flow-on effects in terms of fermentation, wine style, and quality. The regular presence of soil and bark species in ferments has commercial implications as the composition of microbial species present in juice is known to significantly influence the sensory properties of wine (Ciani et al., 2010; Jolly et al., 2014). As well as influencing the physiology of grapevines, vineyard soil and bark also represent important management targets for influencing the types and abundances of fungi that appear in tanks after harvest - particularly for practitioners of “spontaneous” ferments.

Vines are not native to NZ, and viticulture has only been conducted in NZ over the last 100 years. It has been shown that at least one important member of this community, *S. cerevisiae*, has relatively recently been introduced to NZ, probably by humans along with vines (Gayevskiy et al. 2016). The degree to which these observations of connections and origins of ferment communities will hold in other wine producing areas remains to be tested, and the history of viticulture may affect the types and abundance of fungi that are established in vineyards and surrounding unmanaged ecosystems.

Our study not only addresses the origins of wine microbes, but also advances our understanding of complex microbial landscapes and elucidates the nature and connection of various natural microbial communities in a landscape. Further, it also indicates how agricultural and natural ecosystems are interrelated, and this brings into focus the need to consider these microbiomes when sustainability and environmental protection management interventions are developed. Lastly, increasing our understanding of this cryptic microbial landscape may lead to improved management techniques, and open the door to harness and engineer complex microbial ecosystems for more sustainable use.

Experimental Procedures

Sampling

Twenty collaborating companies employing conventional viticultural and winemaking practices provided access to 36 Sauvignon blanc vineyards dispersed across six major wine growing regions in New Zealand: Hawkes Bay, Martinborough, Nelson, Wairau Valley,

Awatere Valley, and Central Otago, and the resulting commercially gathered juice. The maximum geographic distance between any two sampled vineyards spanned 38°- 45° S and 168° - 177° E. The minimum distance between samples of different regions is ~15kms and the maximum ~885kms; the average distance between vineyards within regions is ~15kms. All vineyards and wineries are commercially operated.

Viticultural habitats

Approximately two weeks before harvest, soil, vine bark, and ripe fruit were aseptically collected from each vineyard. Three sub-samples were taken across each block and pooled for each niche, and the pooled sample was ~30g. All samples were taken at least 5m from the edge of vineyards. Soil samples were taken from the topsoil at a depth of < 5cm, 50cm away from a grapevine trunk. Bark samples were taken at least 30cm up the trunk, adjacent to the soil sample. Whole bunches of fruit were aseptically removed and placed into sterile bags. All samples were immediately transported to the laboratory on ice for processing. Epiphytic fungi were washed off fruit samples by being immersed in sterile water and rocked for 30 minutes. The resulting wash was then collected and pelleted down by centrifugation at 3000rpm. The pellet was re-suspended in 500µl of sterile water. Soil and bark samples were homogenised mechanically using aseptic technique to increase surface area for DNA extraction. Before DNA extraction all samples were stored at -20°C. In total 108 vineyard samples were collected and processed.

Native forests

Six samples of each of soil and fruit from different species were similarly sampled aseptically from local native New Zealand conservation reserves in each of the six regions. On average, these reserves were ~20kms from the vineyard samples. Sampled species were those that were fruiting at the time of harvest, and the species include: *Coprosma sp.*, *Coriaria arborea*, *Pittosporum tenuifolium*, *Melicytus ramiflorus*, *Ripogonum scandens*. In total, we collected 36 fruit samples and 36 soil samples from native ecosystems across New Zealand.

Juice and ferments

All fruit was commercially harvested and processed separately. Ten litres of each juice was transferred from the settling tanks in the wineries into sterile containers, and immediately transported on ice to our laboratory. Upon arrival to our facility, each juice was then immediately stirred for 5 minutes, and 50ml of juice was sampled. Containers were fitted with an air lock, and juices were left to spontaneously ferment at 15°C, and after five days another 50ml of a ferment sample was taken. All samples were centrifuged at 3000 rpm and the resulting pellets resuspended in 500 µl of sterile water and frozen at -20°C prior to DNA extraction.

DNA extraction and sequencing

DNA was extracted from all samples using the Zymo Research Soil Microbe DNA MiniPrep™ kits. Fungal communities were characterised and enumerated by 454-sequencing of the D1/D2 region of 26S ribosomal RNA. This locus was amplified by PCR directly from whole

DNA extractions using NL1 and NL4 primers described in Kurtzman and Robnett (2003). Sequencing this locus provides an effective method for quantification and identification of fungal communities (Taylor et al., 2014). Individual samples were amplified with primers tagged with unique multiplex identifiers, to allow samples to be separated bioinformatically. One juice, two fruit, and 17 native samples failed to amplify, and so their respective sample sizes reduced to 35, 34, and 19 respectively. All PCR products were cleaned using AmpureXP beads and their quality checked by Agilent DNA1000 chips. Juice samples were uni-directionally sequenced on a 454-junior instrument by New Zealand Genomics Limited. All remaining samples were sequenced on a full plate of a 454 Life Sciences GS-FLX instrument by Macrogen (Korea). The raw sequences for all samples are present in GenBank (accession number: SRP116986).

Bioinformatic processing pipeline

Sequence processing was carried out using Mothur v.1.30 (Schloss et al., 2009). Primers and sequences <200bp were removed. Low quality reads, including homopolymer errors, were removed using the pyronoise algorithm. Chimeric sequences produced during PCR were identified and removed using the uchime algorithm. The remaining high-quality sequences were bioinformatically assigned unique labels, merged to form one data set, and analysed. Unique sequences in this dataset were compared to SILVA reference database (Quast et al., 2013) of fungal sequences. Sequences that were not identified as fungal were removed (6.4% of all unique sequences, 4.9% of all reads). The remaining unique fungal sequences were then aligned using a freely available fungal reference database and clustered into groups that contained >98% identity. The 98% identity threshold has been empirically determined to sufficiently approximate differences between many fungal species (Kurtzman

and Robnett, 2003; Romanelli et al., 2010). Such >98% identity groups are commonly referred to as molecular operational taxonomic unit (MOTU) and approximate species. Any species represented by a single read was conservatively removed from the dataset. Lastly, before analyses, we randomly sub-sampled (rarefied) DNA sequences from each community to match the sample with the least number of sequences: this produces a data set with an equivalent sampling effort for all communities for all samples. The final dataset totalled 88,200 sequences equally distributed across samples. A representative sequence was then determined for each species as the one with the greatest similarity to all others in that species. Representative sequences were then compared to the fungal Ribosomal Database Project (RPD) reference large subunit rDNA database using a Bayesian approach using the "classify.seqs" command in Mothur (Schloss et al., 2009). This database is able to classify to genus, and all taxonomic levels above. Any sequences that matched 70% or less to a taxonomic classification with the Bayesian approach were listed as unclassified at that particular taxonomic level. The read depth of fungal taxa present across all samples is shown in Table S2.

Statistical analysis

We tested whether the total number of species significantly differs by habitat and/or region by using two-way ANOVA for communities isolated from juice and ferment samples from six different regions. One-way ANOVA tests on species richness were carried out to test the effect of region separately for juice and ferment communities. The differential presence or absence of fungal species was analysed with a two-way full factorial permutational multivariate ANOVA (permanova) (Anderson, 2001) on binary Jaccard dissimilarities, with reported R^2 indicating the amount of variation explained by the factor. One-way tests were

conducted to quantify the effect of region on communities separately. An additive diversity partitioning hierarchical null model test (Crist et al., 2003) for juice and ferment species separately tested whether the distribution of ferment species was randomly distributed across the six regions. Differences in relative abundances of species was tested with a two-way full factorial permutational multivariate ANOVA (permanova) (Anderson, 2001) on standard Jaccard dissimilarities. All statistical testing was conducted using the vegan package (Dixon, 2003) in R (R Core Team, 2016). Differences in community structure across samples were visualised using classical multidimensional scaling of Jaccard dissimilarities. Differences in the overall structure of juice and ferment communities were visualised using Cytoscape software package (Shannon et al., 2003).

Acknowledgements

We thank Sarah Knight and Soon Lee who assisted in sample collection and processing of samples, Peter Tsai for bioinformatic assistance, and Alexandria Leonard for her assistance in editing the manuscript. This work was funded by grants to MG from the New Zealand Ministry Innovation and Employment, Plant and Food Research Ltd and New Zealand Winegrowers. The completion of this research would not have been possible without the enthusiasm, cooperation and assistance of the many collaborating companies who allowed access to their land: Amisfeild, Ata Rangi, Coal Pit, Constellation, Delegats, Domain Road, Frey Vineyard, Misha's Vineyard, Mt Difficulty, Mt Riley, Neudorf, Palliser, Pernod Ricard, Rippon, Seifried, Te Kairanga, Tohu, Trinity Hill, Villa Maria and Vita Brevis.

References

Anderson, M.J., 2001. A new method for non-parametric multivariate analysis of variance. Austral Ecol. 26, 32–46.

Anfang, N., Brajkovich, M., Goddard, M.R., 2009. Co-fermentation with *Pichia kluyveri* increases varietal thiol concentrations in sauvignon blanc. Aust. J. Grape Wine Res. 15, 1–8.

Barata, A., González, S., Malfeito-Ferreira, M., Querol, A., Loureiro, V., 2008. Sour rot-damaged grapes are sources of wine spoilage yeasts. FEMS Yeast Res. 8, 1008–1017.

Barata, A., Malfeito-Ferreira, M., Loureiro, V., 2012. The microbial ecology of wine grape berries. Int. J. Food Microbiol. 153, 243–259.

Blanco, P., Orriols, I., Losada, A., 2011. Survival of commercial yeasts in the winery environment and their prevalence during spontaneous fermentations. J. Ind. Microbiol. Biotechnol. 38, 235–239.

Bokulich, N.A., Collins, T.S., Masarweh, C., Allen, G., Heymann, H., Ebeler, S.E., Millsa, D.A., 2016. Associations among wine grape microbiome, metabolome, and fermentation behavior suggest microbial contribution to regional wine characteristics. mBio 7.

Bokulich, N.A., Ohta, M., Richardson, P.M., Mills, D.A., 2013. Monitoring Seasonal Changes in Winery-Resident Microbiota. PLoS ONE 8(6): e66437

Bokulich, N.A., Thorngate, J.H., Richardson, P.M., Mills, D.A., 2014. Microbial biogeography of wine grapes is conditioned by cultivar, vintage, and climate. Proc. Natl. Acad. Sci. U. S. A. 111, E139–E148.

Börlin, M., Venet, P., Claisse, O., Salin, F., Legras, J.-L., Masneuf-Pomarede, I., 2016. Cellar-associated *Saccharomyces cerevisiae* population structure revealed high-level diversity and perennial persistence at Sauternes wine estates. Appl. Environ. Microbiol. 82, 2909–2918.

Buser, C.C., Newcomb, R.D., Gaskett, A.C., Goddard, M.R., 2014. Niche construction initiates the evolution of mutualistic interactions. Ecol. Lett. 17, 1257–1264.

Ciani, M., Comitini, F., Mannazzu, I., Domizio, P., 2010. Controlled mixed culture fermentation: A new perspective on the use of non-*Saccharomyces* yeasts in winemaking. FEMS Yeast Res. 10, 123–133.

Ciani, M., Fatichenti, F., Mannazzu, I., 2002. Yeasts in winemaking biotechnology. Biodivers. Biotechnol. Wine Yeasts 112–123.

Ciani, M., Mannazzu, I., Marinangeli, P., Clementi, F., Martini, A., 2004. Contribution of winery-resident *Saccharomyces cerevisiae* strains to spontaneous grape must fermentation. Antonie Van Leeuwenhoek Int. J. Gen. Mol. Microbiol. 85, 159–164.

Crist, T.O., Veech, J.A., Gering, J.C., Summerville, K.S., 2003. Partitioning Species Diversity across Landscapes and Regions: A Hierarchical Analysis of α, β, and γ Diversity. Am. Nat. 162, 734–743.

Dixon, P., 2003. VEGAN, a package of R functions for community ecology. J. Veg. Sci. 14, 927–930.

Gayevskiy, V., Goddard, M.R., 2011. Geographic delineations of yeast communities and populations associated with vines and wines in New Zealand. ISME J 6, 1281–1290.

Gayevskiy, V., Lee, S., Goddard, M.R., 2016. European derived *Saccharomyces cerevisiae* colonisation of New Zealand vineyards aided by humans. FEMS Yeast. Res. 16, fow091

Goddard, M.R., 2008. Quantifying the complexities of *Saccharomyces cerevisiae*'s ecosystem engineering via fermentation. Ecology 89, 2077–2082.

Goddard, M.R., Anfang, N., Tang, R., Gardner, R.C., Jun, C., 2010. A distinct population of *Saccharomyces cerevisiae* in New Zealand: Evidence for local dispersal by insects and human-aided global dispersal in oak barrels. Environ. Microbiol. 12, 63–73.

Goddard, M.R., Greig, D., 2015. *Saccharomyces cerevisiae*: A nomadic yeast with no niche? FEMS Yeast Res. 15, fov009

Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., Martiny, J.B.H., 2012. Beyond biogeographic patterns: Processes shaping the microbial landscape. Nat. Rev. Microbiol. 10, 497–506.

Jolly, N.P., Varela, C., Pretorius, I.S., 2014. Not your ordinary yeast: Non-Saccharomyces yeasts in wine production uncovered. FEMS Yeast Res. 14, 215–237.

Knight, M.R., Klaere, S.,Fedrizzi, B., Goddard, M R. 2015. Regional microbial signatures positively correlate with differential wine phenotypes: Evidence for a microbial aspect to terroir. Sci. Rep. 5 :14233.

Knight, S., Goddard, M.R., 2014. Quantifying separation and similarity in a *Saccharomyces cerevisiae* metapopulation. ISME J 9, 361–370.

Knight, S.J., Goddard, M.R., 2015. Sporulation in soil as an overwinter survival strategy in *Saccharomyces cerevisiae*. FEMS Yeast Res. 16, fov102.

Kurtzman, C.P., Robnett, C.J., 2003. Phylogenetic relationships among yeasts of the "*Saccharomyces* complex" determined from multigene sequence analyses. FEMS Yeast Res. 3, 417–432.

Liti, G., Carter, D.M., Moses, A.M., et al. 2009. Population genomics of domestic and wild yeasts. Nature 458, 337–341.

Martiny, J.B.H., Bohannan, B.J.M., Brown, J.H., et al 2006. Microbial biogeography: Putting microorganisms on the map. Nat. Rev. Microbiol. 4, 102–112.

Mercado, L., Dalcero, A., Masuelli, R., Combina, M., 2007. Diversity of *Saccharomyces* strains on grapes and winery surfaces: Analysis of their contribution to fermentative flora of Malbec wine from Mendoza (Argentina) during two consecutive years. Food Microbiol. 24, 403–412.

Morrison-Whittle, P., Goddard, M.R., 2015. Quantifying the relative roles of selective and neutral processes in defining eukaryotic microbial communities. ISME J. 9, 2003–2011.

Mortimer, R., Polsinelli, M., 1999. On the origins of wine yeast. Res. Microbiol. 150, 199–204.

Nemergut, D.R., Schmidt, S.K., Fukami, T., O'Neill, S.P., Bilinski, T.M., Stanish, L.F., Knelman, J.E., Darcy, J.L., Lynch, R.C., Wickey, P., Ferrenberg, S., 2013. Patterns and processes of microbial community assembly. Microbiol. Mol. Biol. Rev. 77, 342–356.

Pinto, C., Pinho, D., Cardoso, R., Custódio, V., Fernandes, J., Sousa, S., Pinheiro, M., Egas, C., Gomes, A.C., 2015. Wine fermentation microbiome: A landscape from different Portuguese wine appellations. Front. Microbiol. 6, 905

Pretorius, I.S., 2000. Tailoring wine yeast for the new millennium: Novel approaches to the ancient art of winemaking. Yeast 16, 675–729.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D59

R Core Team, 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Romanelli, A.M., Sutton, D.A., Thompson, E.H., Rinaldi, M.G., Wickes, B.L., 2010. Sequence-based identification of filamentous basidiomycetous fungi from clinical specimens: A cautionary note. J. Clin. Microbiol. 48, 741–752.

Santamaría, P., López, R., López, E., Garijo, P., Gutiérrez, A.R., 2008. Permanence of yeast inocula in the winery ecosystem and presence in spontaneous fermentations. Eur. Food Res. Technol. 227, 1563–1567. doi:10.1007/s00217-008-0855-5

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Stefanini, I., Dapporto, L., Legras, J.-L., Calabretta, A., Di Paola, M., De Filippo, C., Viola, R., Capretti, P., Polsinelli, M., Turillazzi, S., Cavalieri, D., 2012. Role of social wasps in *Saccharomyces cerevisiae* ecology and evolution. Proc. Natl. Acad. Sci. U. S. A. 109, 13398–13403.

Taylor, M.W., Tsai, P., Anfang, N., Ross, H.A., Goddard, M.R., 2014. Pyrosequencing reveals regional differences in fruit-associated fungal communities. Environ. Microbiol. 16, 2848–2858.

Zhang, H., Skelton, A., Gardner, R.C., Goddard, M.R., 2010. *Saccharomyces paradoxus* and *Saccharomyces cerevisiae* reside on oak trees in New Zealand: Evidence for migration from Europe and interspecies hybrids. FEMS Yeast Res. 10, 941–947.

Zott, K., Claisse, O., Lucas, P., Coulon, J., Lonvaud-Funel, A., Masneuf-Pomarede, I., 2010. Characterization of the yeast ecosystem in grape must and wine using real-time PCR. Food Microbiol. 27, 559–567.

Tables

Table 1: Results of additive diversity and null model testing of juice (n=35) and ferment (n=33) communities (9999 permutations).

		α_1	α_2	β_1	β_2	γ
Juice	Observed	16.89	47.17	30.28	79.83	127
	Simulated	32.44	67.71	35.28	59.29	127
	P-values	0.0001	0.0001	0.0001	0.0001	1
Ferment	Observed	4.85	16.33	11.48	39.67	56
	Simulated	11.73	23.85	12.12	32.15	56
	P-values	0.0001	0.0001	0.1553	0.0001	1

Figure legends

Figure 1: The three measures of community diversity examined.

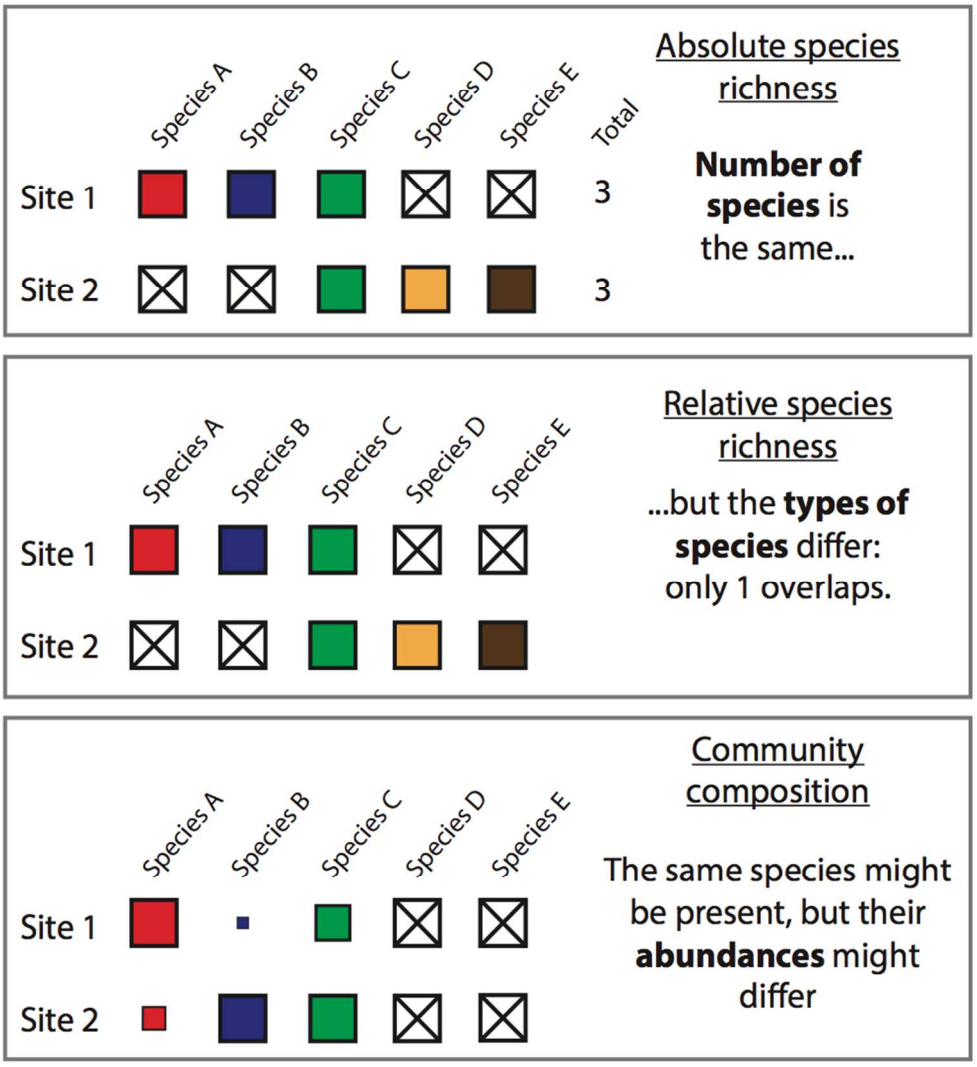
Figure 2: The relative abundance of the most abundant fungal species (OTUs comprising >1% of all juice or ferment reads) found in juice or ferment samples as well as their relative abundance across the six wine growing regions.

Figure 3: Similarity of community composition (Jaccard similarity) of juice and ferment samples when compared to samples from the same region as well as to samples from different regions.

Figure 4: Number of species in native forests (43 samples), vineyards (108), juice (35), and ferment (33) communities, and the number of species that connect each compartment. Note, the overlap is calculated separately for each comparison.

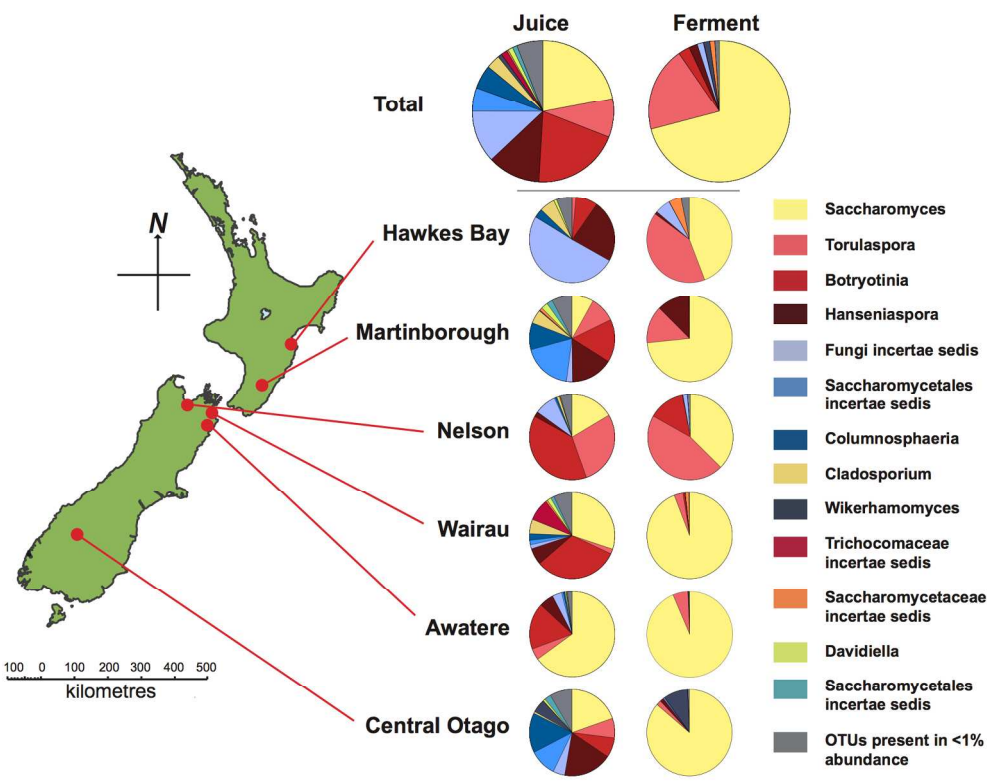
Figure 5: A) Relative abundance of fungal species present in both juice (blue circles) and ferment (red circles) where node size is proportional to relative abundance. B) Relative frequency of ferment species in vineyard habitats.

Figure 6: Ternary plot of juice and ferment communities and their % overlap with bark, fruit, and soil vineyard communities isolated from the vineyard from which the juice/ferment samples were harvested.



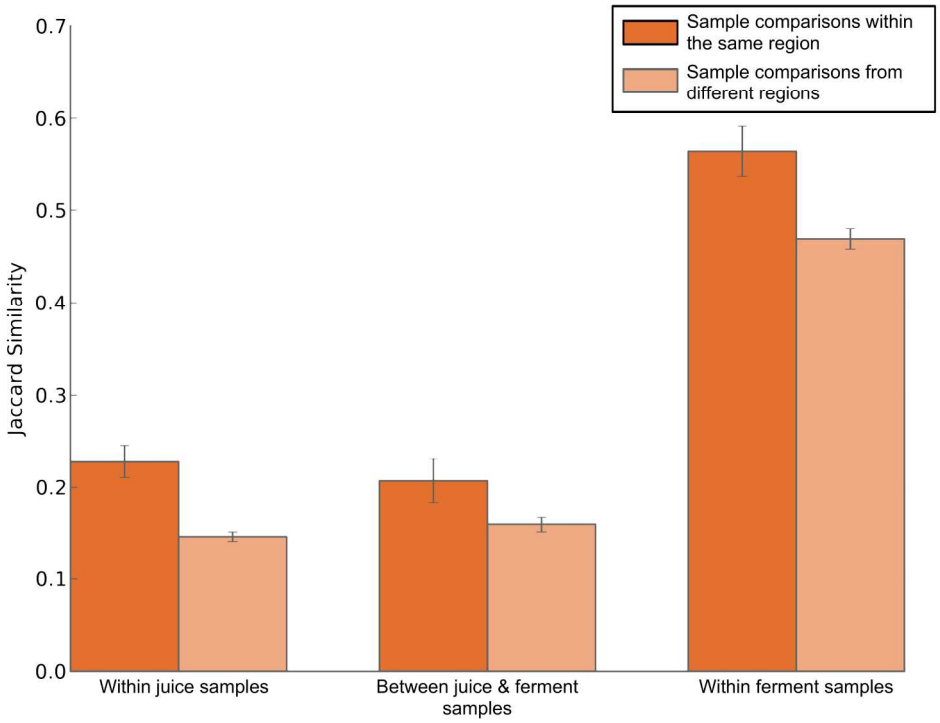
The three measures of community diversity examined.

86x92mm (300 x 300 DPI)



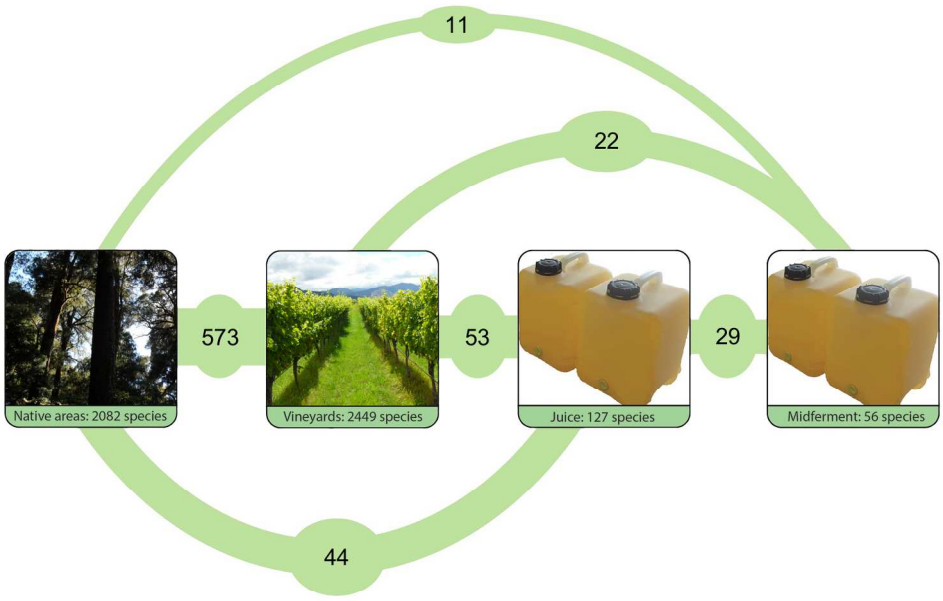
The relative abundance of the most abundant fungal species (OTUs comprising >1% of all juice or ferment reads) found in juice or ferment samples as well as their relative abundance across the six wine growing regions.

177x138mm (300 x 300 DPI)



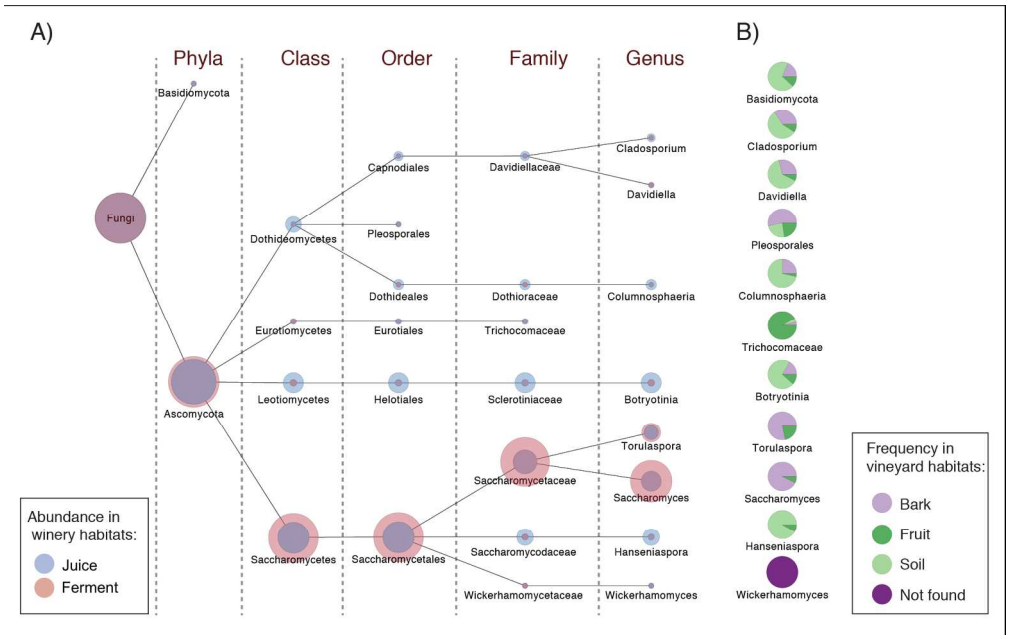
Similarity of community composition (Jaccard similarity) of juice and ferment samples when compared to samples from the same region as well as to samples from different regions.

195x146mm (300 x 300 DPI)



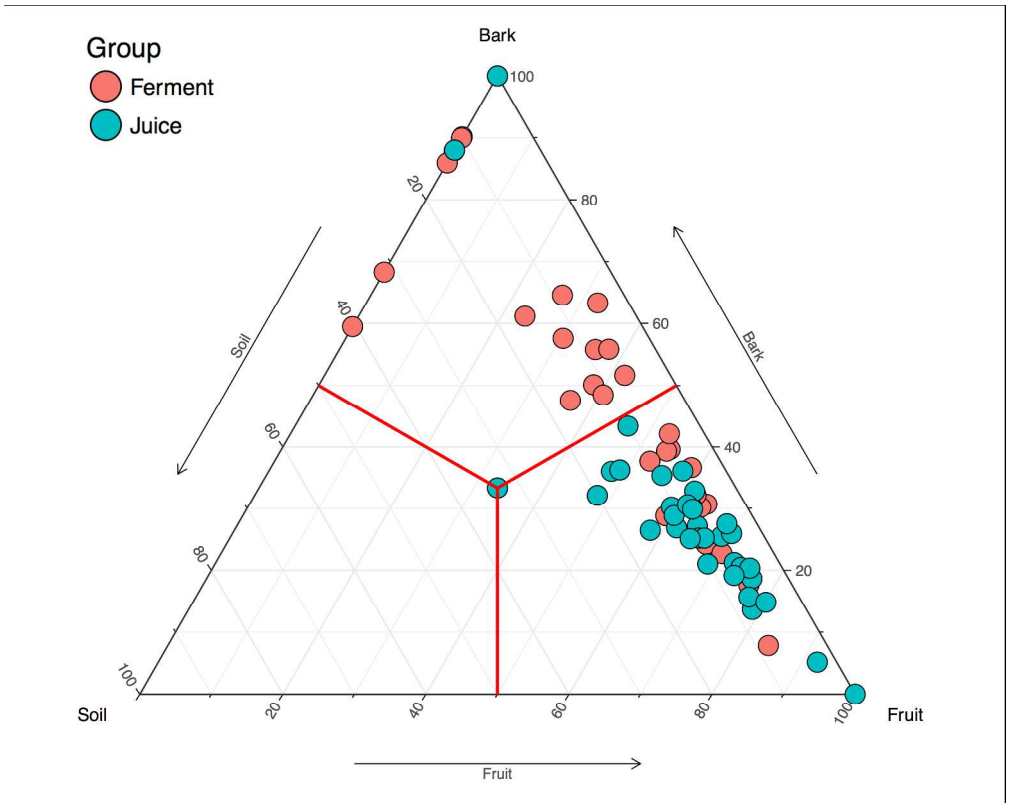
Number of species in native forests (43 samples), vineyards (108), juice (35), and ferment (33) communities, and the number of species that connect each compartment. Note, the overlap is calculated separately for each comparison.

177x125mm (300 x 300 DPI)



A) Relative abundance of fungal species present in both juice (blue circles) and ferment (red circles) where node size is proportional to relative abundance. B) Relative frequency of ferment species in vineyard habitats.

177x112mm (300 x 300 DPI)



Ternary plot of juice and ferment communities and their % overlap with bark, fruit, and soil vineyard communities isolated from the vineyard from which the juice/ferment samples were harvested.

237x189mm (300 x 300 DPI)